Structure and coding properties of a dominant *Escherichia coli* mutator gene, *mutD*

(proofreading/DNA replication/Bal-31 nuclease deletion mapping)

Edward C. Cox and Deborah L. Horner

Department of Biology, Princeton University, Princeton, New Jersey 08544

Communicated by J. T. Bonner, December 30, 1982

ABSTRACT The region of the *Escherichia coli* chromosome coding for the *mutD* gene was cloned. *mutD5* function resides on a 1.2-kilobase fragment coding for a 28-kilodalton (kDa) protein. A deletion end-point analysis shows that the presence of the 28-kDa protein is required for *mutD5* function and suggests that the *mutD* functional region has sufficient capacity to code for a second polypeptide of approximately 20 kDa. Plasmids carrying the *mutD5* and *mut*⁺ alleles both produce the 28-kDa species. The product of *mutD5* is dominant when carried by single and multicopy plasmids. The product of *mutD5* exhibits negative complementation. We suggest that the 28-kDa protein participates in a multimeric structure, perhaps at the replication fork.

mutD is a dominant Escherichia coli mutator whose mutation rate is under the control of thymidylate pools (1). All classes of transition and transversion mutation rates and many frameshift mutation rates are increased from 50- to 10,000-fold when mutD strains are grown in rich broth or minimal medium supplemented with thymidine (2), whereas mutD strains grown in minimal medium exhibit mutation rates only 10- to 100-fold higher than wild type. Six mutD alleles have been isolated and studied (3). All are dominant, noncomplementing in trans, and mutate at the same rate when grown in broth. Two, however, do not respond to added thymidine. The function of the mutD gene is otherwise unknown.

In this report we assign *mutD* function to a region of the genome approximately 1.2 kilobase pairs (kb) long. By comparing the proteins encoded by this region to a set of exonuclease-constructed deletions, we show that the 1.2-kb segment codes for at least one protein of 28 kilodaltons (kDa). We also show that *mutD5* exhibits negative complementation, acting as a dominant gene when in the presence of a single *mut*⁺ copy but behaving as an incompletely recessive gene in the presence of a multicopy *mut*⁺ plasmid.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. KH1213 (mut^+) and KH1214 (mutD5) are *E. coli* K-12 derivatives of W3110 (3). JC1569 is $recA^-$ and is described in ref. 4. CSR603 is an ultraviolet-sensitive strain used in the maxicell method of Sancar *et al.* (5). 803 is *supF recA met^-* and was obtained from K. Murray. pBR322 and pACYC177 were received from A. C. Y. Chang (4).

Enzymes and Biochemicals. Restriction endonucleases, T4 ligase, and the exonuclease Bal-31 were purchased from Bethesda Research Laboratories and New England BioLabs. They were used according to the manufacturers' recommendations. [^{35}S]Methionine, approximately 1,000 Ci/mmol (1 Ci = 3.7 × 10¹⁰ becquerels), was purchased from Amersham.

Cloning Strategy. A Sau3A partial digest of KH1214 (mutD5) DNA was sized on a sucrose density gradient. A fraction of approximately 5.0 kb was ligated into the BamHI site of pACYC177, and kanamycin-resistant transformants of JC1569 were selected, purified, and grown to saturation in L-broth (2). Individual cultures were then tested for mutator activity (3). Five mutator-active (Mut⁻) clones were isolated by screening approximately 500 isolates. One, pREC31, was purified (6), digested with EcoRI, and religated. The religated product, pDH1, was digested with EcoRI and HindIII and subcloned into pBR322 that had been digested with EcoRI and HindIII. (The structure of one such subclone, pEC66, is shown in Figs. 1 and 3.) With knowledge of the restriction map in hand, pEC66 was used as a probe to identify a 1.6-kb mut⁺ fragment from an EcoRI digest of KH1213 (mut⁺). The digest was separated on an agarose gel and probed with nick-translated (7) pEC66 by the method of Southern (8). The 1.6-kb region of a preparative agarose gel was then isolated and ligated into the EcoRI site of the M13 phage vector mp8 (9). Subclones carrying regions homologous to pEC66 were recognized by probing M13 plaques with nicktranslated pEC66. Replicative intermediates from one isolate were cloned into pBR322 to yield pDH105, which was used to obtain the results shown in Table 1 and Fig. 3.

Exonuclease Bal-31 Deletions. Two pEC66 preparations were separately cut with *Eco*RI and *Hin*dIII and digested for various lengths of time with sufficient Bal-31 to remove from a very few to several thousand base pairs (bp) from both ends of the plasmid DNA. A typical reaction contained 20 μ g of cut plasmid DNA in a 200- μ l reaction mixture containing 2.5 units of Bal-31. Digestion was carried out at 30°C for a total of 180 min. Samples were removed at 30, 60, 120, and 180 min, extracted twice with buffered phenol, precipitated with two vol of ethanol, resuspended in 1 mM Tris⁺HCl, pH 8.0/0.1 mM EDTA, and ligated with T4 ligase (see Fig. 1). These preparations were used to transform 803. Transformants from each of the 30-, 60-, 120-, and 180-min Bal-31 digests were checked for mutator activity. From this set a subset was chosen for further analysis (see Fig. 2).

Labeling and Identifying Plasmid Protein Products. Various pACYC177 and pBR322 derivatives were used to label plasmidcoded proteins by the maxicell system described by Sancar *et al.* (5) modified to include a cycloserine selection step (10). Labeled cells were lysed and run on NaDodSO₄/polyacrylamide gels as described by Laemmli (11). Fixed and dried gels were exposed at -70° C by the procedure of Bonner and Lasky (12) with Kodak X-Omat AR film.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: kb, kilobase pairs; bp, base pair(s); kDa, kilodaltons; Mut⁻, mutator-active phenotype; Mut⁺, wild-type phenotype; CCCP, carbonylcyanide *m*-chlorophenylhydrazone.

RESULTS

mutD5 and mut⁺ Clones. Several Mut⁻ transformants were recovered as inserts in pACYC177, confirming that mutD5 is dominant (3). One, pREC31, was purified, digested with EcoRI, and religated. A small plasmid preparation (13) isolated from a Mut⁻ transformant was then digested with EcoRI and HindIII and cloned into an EcoRI/HindIII digest of pBR322 to yield a Mut⁻ ampicillin-resistant plasmid, pEC66. It contains a 1.6-kb insert with a Mut⁻ fragment from KH1214 joined at the original Sau3A/BamHI junction to a 560-bp pACYC177 fragment (Figs. 1 and 2). With partial knowledge of the pEC66 restriction map available, a 1.6-kb fragment from an EcoRI digest of KH1213 (mut^+) DNA was first cloned in bacteriophage M13 mp8 and then subcloned into pBR322 as described. A restriction analysis of this clone (pDH105) with BamHI, EcoRI, Acc I, and HinfI gave a restriction map (data not shown) identical to that for the mutD region of pEC66 shown in Fig. 2. Apparently the EcoRI site at 1.6 kb on pDH105 is just to the right of the Sau3A/BamHI junction of pEC66.

The mutation frequencies of various strains carrying mut^+ and mutD5 on pBR322 are summarized in Table 1. These frequencies were measured with strains grown in L-broth, where the conditional mutator phenotype of mutD is maximally expressed (14). With mutD5 in trans, the mutation rate is elevated several orders of magnitude, regardless of the allele on the chromosome. However, when mut^+ is on the plasmid and mutD5 on the chromosome, mutation frequencies were depressed to within 1 order of magnitude of a haploid wild-type culture. Thus, mutD5 exhibits an almost completely recessive phenotype on the chromosome in trans with mut^+ on a multicopy plasmid. Many copies of the mut^+ gene neither depress nor elevate the mutation rate (Table 1, line 3).

Deletion Analysis. Bal-31-generated deletions extending from the *Eco*RI site and from the *Hin*dIII site were isolated as described in the legend to Fig. 1. Each deletion was then assayed for mutator activity, and a subset was used to transform CSR603 for an analysis of gene products (Figs. 2–4).

The general result is that deletion end points from both ends of the 2.1-kb insert of pEC66 define a region of mutator activity that is 1.1–1.2 kb long. The deletions extending from the *Hin*dIII site (Fig. 2) can be divided unambiguously between mutator and nonmutator clones, placing the right end of the mutator function between the 1.3-kb *Bam*HI site and the *Acc* I site. The left end falls between the pDH45 and pDH53 endpoints and, thus, lies within approximately 100 bp of the *Eco*RI site.

mutD and mut⁺ Both Code for a 28-kDa Protein. [³⁵S]Methionine-labeled samples from CSR603 strains with and without various plasmids were run on polyacrylamide/NaDodSO₄ gels (Fig. 3). Because the 28-kDa protein runs quite close to the 29-kDa β -lactamase coded for by pBR322, samples were labeled in the presence and absence of an inhibitor of membrane polarity. In the presence of CCCP, the precursor of β -lactamase is not processed (15) and separates convincingly from the 28-kDa species.

The results summarized in Fig. 3 show that the 28-kDa protein is present in CSR603 strains carrying both the mut^+ and mutD5 alleles. Sample wells loaded with CSR603 and CSR603 (pBR322) occasionally reveal a protein slightly smaller than 28 kDa (see, for example, Fig. 3, lanes c and d), but this protein is always faintly labeled and clearly smaller than the 28-kDa species.

mutD5 Function Is Correlated with the Presence of the 28kDa Protein. pDH deletions were examined by the maxicell labeling procedure for the presence or absence of the 28-kDa protein. The general result is that all Mut⁺ deletions extending from the HindIII site lack the 28-kDa protein, whereas those from the EcoRI site show variable behavior. The data for the deletions that help define the left and right boundaries of the mutator function are shown in Fig. 4. All of the deletions extending from the HindIII site that are Mut⁻ also code for the 28-kDa protein. Some have lost the 1.3-kb BamHI site (pDH24, 26, 25), but all retain the Acc I site. The deletions that have lost mutator activity invariably lack the 28-kDa protein. One, pDH10, lacks the 28-kDa protein and retains the Acc I site. Thus, the right-hand terminus of mutator function must be between 1.1 and 1.3 kb on our map, consistent with the inference drawn from the deletion set that the boundary lies between pDH6 and pDH10 at 1.2 kb.

Deletions extending from the *Eco*RI site display a variety of altered gel patterns. Both Mut⁻ deletions shown in Fig. 4, pDH49 and pDH45, code for the 28-kDa protein. Two Mut⁺ deletions appear to produce altered 28-kDa polypeptides,



FIG. 1. Construction of deletions extending through the *mutD* region. pEC66 was isolated as described. Deletions from the *Hind*III site, shown above, were constructed by incubating *Hind*III-digested pEC66 with exonuclease Bal-31 for various lengths of time, followed by ring closure with T4 ligase. Deletions from the *Eco*RI site were constructed by the same procedure with an *Eco*RI-digested sample of pEC66. The *Hind*III site is in the *tet* promoter; *tet* and *amp* are the genes for tetracycline resistance and ampicillin resistance, respectively.



FIG. 2. The *mutD5* region of pEC66. pDH subclones of pEC66 were purified, and small plasmid preparations (13) were used to transform CSR603 for the maxicell analysis summarized in Fig. 4. Each preparation was cut with *Pst* I at the unique *Pst* I site in the *amp* gene (see Fig. 1) and was sized by comparing the mobility on an agarose gel of the cut DNA to an *EcoRI/Hind*III double digest of phage λ DNA (4). The 1.6-kb fragment (m) came from KH1214 (*mutD5*), whereas the DNA between the *Sau3A/Bam*HI junction and the *Hind*III site (m) is from pACYC177 (see Fig. 1). The length of each deletion, represented as a heavy line, was estimated as half of the total amount of DNA removed by Bal-31 digestion (see *Discussion*). MutD5 means that the dominant mutator phenotype is retained in a deletion; Mut⁺ (wild type) means that it has been lost. pDH27 and pDH6 have the 1.3-kb *Bam*HI site; all other deletions longer than pDH27 extending from the *Hind*III site lack it. pDH clones 24, 26, 25, and 10 retain the *Acc* I site, whereas clones 30-36 on the *Hind*III site have lost it.

pDH53, coding for a 27-kDa polypeptide (which, however, may correspond to a faintly labeled protein of this molecular mass occasionally seen in this region of the gel), and pDH55, coding

Table 1. Mutation frequencies in mut⁺ and mutD5 strains

Strain	Genotype		$\begin{array}{r} \text{Mutation frequency} \\ \times \ 10^8 \ \text{to}^* \end{array}$	
	Chromosome	Plasmid	Nal ^R	Rif ^R
KH1213	mut ⁺		0.3	0.8
KH1214	mutD5	_	8,530	15,000
KH1265	mut^+	mut ⁺	0.2	0.9
KH1266	mut^+	mutD5	2,090	15,200
KH1267	mutD5	mut^+	11.2	47.2
KH1268	mutD5	mutD5	5,270 ·	17,500

Nal^R, resistance to nalidixic acid; Rif^R, resistance to rifampicin. * Frequencies of mutation to Nal^R and Rif^R phenotypes are the average of five individual determinations, each started from a mutant-free inoculum. The cells were grown in L-broth. For additional details, see ref. 3. The plasmids carried by KH1265, KH1266, KH1267, and KH1268 were pDH105, pEC66, pDH105, and pEC66, respectively.



FIG. 3. mutD5 and mut⁺ plasmids code for a 28-kDa protein. [³⁵S]-Methionine was used to label cells carrying various plasmids by the procedure of Sancar et al. (10). Labeled cells were lysed, electrophoresed in 12% acrylamide gels, and exposed to x-ray film. Just before the labeling step, each sample was divided in half. One-half was incubated for 5 min with 20 μ M carbonylcyanide m-chlorophenylhydrazone (CCCP) before adding the [³⁵S]methionine. This step prevents processing of the pre- β -lactamase (15) and makes the 30-kDa region of the gel easier to interpret. Gels were calibrated with ¹⁴C-labeled protein standards; the molecular masses of pre- β -lactamase and β -lactamase are 31 kDa and 29 kDa, respectively (16, 17). The calculated molecular mass of the 28kDa species is 27.9 \pm 0.8 (n = 5). Lanes: a, CSR603; b, CSR603 with 20 μ M CCCP; c, CSR603 (pBR322); d, CSR603 (pBR322) with 20 μ M CCCP; g, CSR603 (pDH105 mut⁺); h, CSR603 (pDH105 mut⁺) with 20 μ M CCCP.

for two new polypeptides of about 27 kDa and 26 kDa (not well resolved in Fig. 4). Neither pDH59 nor pDH47, although Mut⁺, seem to code for an altered 28-kDa product, although very small changes in molecular weight cannot be detected in these gels. These two plasmids represent the only cases in which the loss of mutator activity is not paralleled also by an easily detected loss or alteration of the 28-kDa protein.

DISCUSSION

Our results identify a region of approximately 1,200 bp with both *mutD5* function and a 28-kDa protein. The fact that *mutD5* is dominant allowed us to correlate the disappearance of a gene product with a series of roughly ordered deletions. Although it is clear that a hierarchy of deletion end points can readily be constructed (Fig. 2) there is ambiguity in this method because Bal-31 does not necessarily digest both ends of double-stranded DNA at the same rate (18). Nonetheless, with the exceptions of pDH6 and pDH27, the Mut⁻/Mut⁺ borders at both ends of the insert deduced by this method are consistent with the *Bam*HI and *Acc* I restriction map and with the protein labeling patterns. This consistency argues that the deletion size and order inferred by these methods is sufficiently precise to locate the mutator boundaries within 100 bp.

We have presented the data in Fig. 4 in the order consistent with the deletion map shown in Fig. 2. However, the order of pDH53, -59, -55, and -47 is arbitrary because they differ from each other by no more than 50 bp. We cannot reliably make distinctions at this level. The fact that pDH53 may code for a



FIG. 4. Deletions across the *mutD* boundaries. The deletions in Fig. 2 were examined in the maxicell system. Only those crossing the Mut⁻/Mut⁺ boundaries are shown here. The numbers under each lane correspond to pDH clone numbers in Fig. 2. pDH55 codes for two new polypeptides, not clearly resolved here, of ≈ 27 and ≈ 26 kDa. (*Right*) two control lanes, CSR603 (pEC66) and CSR603 (pBR322), respectively Mut⁻ and Mut⁺. CCCP was not used in these experiments, and the acrylamide concentration in the gel was 10%.

smaller polypeptide and that pDH55 clearly does so suggests that the real order is (59,47), 53, and 55.

It is a striking fact that all Mut⁻ deletions from the HindIII site code for the 28-kDa protein. Those from the EcoRI site, however, show more complicated gel patterns. Of the four Mut deletions shown here, pDH53 appears to code for a protein smaller than 28 kDa, pDH55 codes for two proteins of approximately 27 kDa and 26 kDa, and two, pDH59 and pDH47, code for a protein not detectably altered. These results suggest that the promoter for the 28-kDa protein is proximal to the HindIII site and that deletions extending from the EcoRI site code for polypeptides lacking portions of their carboxyl terminus. If this is so, pDH55 must be deleted in such a way that a weak and a strong chain-terminating codon are created in EcoRI-proximal DNA by fusion to the mutD5 insert. This interpretation is consistent with our observation that all deletions extending from the HindIII site show an abrupt transition between the presence and absence of the 28-kDa species because we expect deletions extending into a promoter sequence to abolish transcription. This interpretation of the results is consistent with a recent finding in this laboratory that certain deletions spanning the EcoRI site fuse a portion of the 28-kDa protein to the β -lactamase gene product (unpublished data). Because the β -lactamase gene of pBR322 is transcribed counterclockwise relative to the EcoRI site (19), fusion products are expected only for the *mutD* orientation suggested here.

A protein of 28 kDa requires a coding region of approximately 750 bp. Because the maxicell system and one-dimensional gels cannot be expected consistently to reveal every region coding for a polypeptide, it is always possible that an additional protein of approximately 20 kDa is coded for in the *mutD* region. Careful examination of many gels has not revealed any differences between pBR322 and either mut^+ - or *mutD5*-encoded proteins in the 20-kDa region, however. If such a cistron exists, the coding region would be proximal to the *Hind*III site because it is the 28-kDa protein that is altered by deletions from the *Eco*RI site.

In *mutD* strains, all classes of point mutations and many frameshifts are elevated several orders of magnitude. Mutation is closely tied to DNA replication (20). One explanation proposed for these results is that the editing function of the DNA replication complex is inhibited (20, 21) either (*i*) because *mutD* interferes directly with the structure of the replication fork or (*ii*) because the *mutD* product fails to repress the level of an inhibitor of proofreading (20). Our results do not allow us to distinguish between these two possibilities. We have shown that mut⁺ and mutD5 both code for a 28-kDa protein in the maxicell system. It is also clear that mutD5 is dominant when both mut^+ and mutD5 are present in low copy number (3), yet incompletely recessive when mut⁺ is carried on a multicopy plasmid, thus reinforcing our view that the mutD5 product participates in a multimeric structure. The proposed multimer could be a component of the replication fork, the model favored by Degnen (20), or, by analogy with lacIs mutations (22), part of a regulatory multimer containing the 28-kDa protein. The function of the regulatory multimer would be to repress the synthesis of a proofreading inhibitor. In mutD strains, the inhibitor would be turned on. The most parsimonious explanation for the phenotype characteristic of mutD is clearly (i), and in this regard it is worth noting that the promoter for the 28-kDa protein appears to be relatively strong, almost as strong as the pBR322coded β -lactamase promoter (Fig. 3). This would be a surprising result for a repressor protein. Whatever the detailed mechanism may be, thymidylate pools (1, 3, 14) turn on the high mutation rate in mutD5 strains, and this fact must also be explained by a satisfactory model for *mutD* action.

The phenotype characteristic of *mutD* is similar in several respects to that of a strong mutator isolated by Horiuchi et al. (23) and named dnaQ49. Both mutators map at 5 min on the E. coli chromosome, are sensitive to 5-aminoacridine, and can be suppressed by mutations in the nalA gene (20, 24; see Horiuchi as quoted in ref. 25). However, the dnaQ49 mutation is recessive (23) and exhibits a salt-dependent DNA elongation phenotype at 44.5°C (26). A dnaQ⁺ clone recently has been shown to code for two proteins, a 21-kDa species that copurifies with RNase H and a 25-kDa species that, when inactivated by $\gamma\delta$ insertion, proves to be the $dnaQ^+$ product (25). Given the very similar map position and other phenotypes of mutD and dnaQ, it is possible that the *mutD* and *dnaQ* mutations lie in the same gene. The difference in complementation behavior of the two would be explained if the dnaQ49 product is altered in such a way that it cannot compete successfully with wild-type subunits to form the multimer postulated above.

We are grateful to E. McCauley, R. E. Cox, and K. Young for their help with various aspects of this work. This study was supported by grants from the Whitehall Foundation and a National Institutes of Health Grant GM28923.

- 1. Degnen, G. E. & Cox, E. C. (1974) J. Bacteriol. 117, 477-487.
- Fowler, R. G., Degnen, G. E. & Cox, E. C. (1974) Mol. Gen. Genet. 133, 179–191.
- 3. Cox, E. C. & Horner, D. L. (1982) Genetics 100, 7-18.

- Chang, A. C. Y. & Cohen, S. N. (1978) J. Bacteriol. 134, 1141-4. 1156.
- Sancar, A., Hack, A. M. & Rupp, W. D. (1979) J. Bacteriol. 137, 5. 692 - 693
- Staudenbauer, W. L. (1976) Mol. Gen. Genet. 145, 273-280. 6.
- 7. Rigby, P. W., Dieckmann, M., Rhodes, C. & Berg, P. (1977) J. Mol. Biol. 113, 237-251.
- Southern, E. M. (1975) J. Mol. Biol. 98, 503-517. 8.
- 9
- Messing, J. & Vieira, J. (1982) Gene 19, 269–276. Sancar, A., Wharton, R. P., Seltzer, S., Kacinski, B. M., Clarke, 10. N. D. & Rupp, W. D. (1981) J. Mol. Biol. 148, 45–62. Laemmli, U. K. (1970) Nature (London) 227, 680–685. Bonner, W. M. & Lasky, R. A. (1974) Eur. J. Biochem. 46, 83–88. Klein, R. D., Selsing, E. & Wells, R. D. (1980) Plasmid 3, 88–91.
- 11.
- 12.
- 13.
- Erlich, H. A. & Cox, E. C. (1980) Mol. Gen. Genet. 178, 703-708. 14.
- 15. Daniels, C. J., Bole, D. G., Quay, S. C. & Oxender, D. L. (1981) Proc. Natl. Acad. Sci. USA 79, 5396-5400.
- Ambler, R. P. & Scott, G. K. (1978) Proc. Natl. Acad. Sci. USA 16. 75, 3732-3736.
- Sutcliffe, J. G. (1978) Proc. Natl. Acad. Sci. USA 75, 3737-3741. 17.

- 18. Gray, H. B., Winston, T. P., Hodnett, J. L., Legerski, R. J., Nees, D. W., Chik-Fong, W. & Robberson, D. L. (1981) in Gene Amplification and Analysis, eds. Chirikjian, G. G. & Papas, T. S. (Elsevier/North-Holland, New York), Vol. 2, pp. 169–203.
- Sutcliffe, J. G. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 19. 77-90.
- 20. Degnen, G. E. (1974) Dissertation (Princeton University, Princeton, NJ).
- Topal, M. D. & Fresco, J. R. (1976) Nature (London) 263, 285-21. 289.
- 22. Muller-Hill, B., Crapo, L. & Gilbert, W. (1968) Proc. Natl. Acad. Sci. USA 59, 1259–1264. Horiuchi, T., Maki, H. & Sekiguchi, M. (1978) Mol. Gen. Genet.
- 23. 163, 277-283.
- Cox, E. C. (1976) Annu. Rev. Genet. 10, 135-156. 24.
- 25. Horiuchi, T., Hisaji, M., Maruyama, M. & Sekiguchi, M. (1981) Proc. Natl. Acad. Sci. USA 78, 3770-3774.
- Horiuchi, T., Maki, H. & Sekiguchi, M. (1981) Mol. Gen. Genet. 26. 181, 24-28.